Bayesian Analysis of Power Transformer Failure Rate Based on Condition Monitoring Information

Zhong Zhang
zzkerry@iastate.edu

Yong Jiang
jiangy@iastate.edu

James D. McCalley
jdm@iastate.edu

Iowa State University, Ames IA 50010

Abstract-- **In this paper, we propose a Bayesian approach for estimating the failure rate of power transformers. The key feature of our approach is that now we are able to use, in a formal manner, the condition monitoring information in failure rate estimation as well as maintenance optimization for power transformers. This is accomplished by using likelihood function which is constructed based on available condition monitoring information to update the assumed probabilistic model of transformer failure rate. We illustrate our approach via some real transformer condition data.**

Keywords--**Transmission equipment, condition monitoring, maintenance optimization, power transformer failure rate, Bayesian analysis**.

I. Introduction

The development in sensor, communication, computer, and data storage technologies has allowed the realization of a variety of condition monitoring systems for bulk transmission system equipment, e.g., power transformers, circuit breakers, transmission lines, and other related equipment, in order to utilize these capital-intensive transmission equipment in the optimal manner. The condition data includes inspection data, test and sample data, SCADA/EMS data, as well as real-time monitoring data. An important issue related to the accumulation of so much data is how to most effectively utilize it. Under the continuous pressure to reduce costs and to maintain reliability, there is a growing need for the utility industry to better utilize the equipment condition monitoring information to optimize the maintenance activities. The industry has developed some tools, work management tools such as CASCADE [1], MAXIMO [2], and INDUS [3], and data integrators such as e.g., Maintenance Management Workstation from EPRI [4], and Asset Sentry from ABB [5]. These systems are, taken together, good at collecting, storing, integrating, and displaying data; however, the coded "intelligence" to use the data for decision-making has not yet evolved.

In this paper, we describe a Bayesian approach to estimate the equipment's failure probability, which enables the development of a risk-based maintenance optimization system for transmission equipment [6]. This framework provides the ability to select and schedule maintenance tasks so as to utilize the available financial and human resources to optimize the risk-reduction achieved from them within a given budget cycle. It is important to observe the significance of this objective in that it differs from the traditional utility objective of minimizing costs subject to some constraint on minimal maintenance achievement. It also differs from the long-term objective of maximizing equipment life.

Although the Bayesian approach proposed in this paper is applicable to all transmission system equipment, in order to limit our discussion, we focus on power transformers.

This paper is organized as follows: Section II presents the overall maintenance optimization problem for transmission system. Section III introduces some fundamental concepts of equipment reliability. Section IV describes the Bayesian approach of estimating power transformer failure rate based on available condition monitoring data. Section V gives an illustrative example of this approach. Section VI concludes.

II. Maintenance Optimization Problem

Equipment maintenance is costly but essential to ensuring transmission system reliability. Its effectiveness can vary dramatically depending on the target and timing of the maintenance activities. The state of the art maintenance practice in the industry goes by the term "reliability centered maintenance" (RCM), which prioritizes maintenance activities based on quantification of likelihood and consequence of equipment failures.

The objective of the work reported in this paper is to develop a method of maximizing effectiveness of the maintenance activities (i.e. maximizing the cumulative system risk reduction) subject to constraints on economic resources, available maintenance crews, and restricted time intervals. We have developed an optimization problem and solution algorithm, described in [6]. Here we just provide the attributes of this problem that are germane to the objective of this paper. The risk for a particular contingency k at time t is the expectation of severity associated with that contingency and loading conditions at that time, computed by summing over possible security-related outcomes (overload, cascading overloads, low voltage, voltage collapse) the product of the outcome probability and its severity [7]. The cumulative risk for that contingency, $CR(k)$, is obtained by accumulating the risk for that contingency over time. We have developed a simulator to perform this calculation. Given the cumulative risk for a contingency k together with its probability $p(k)$, and a maintenance task m that reduces the contingency probability by $\Delta p(k,m)$, we can get the risk reduction associated with that maintenance task performed at time t as $\Delta CR(k,m,t) = \{\Delta p(k,m)/p(k)\} * CR(k)$. Let N be the number of maintainable transmission components and L_k be the number of maintenance *levels* for component k. Let $k = 1,\ldots,N$ be the index over the set of transmission components, $m = 1,\ldots,L_k$ be the index over the set of maintenance activities for transmission component *k*, and $t = 1,\ldots,T$ be the index over the time periods. Define $Iselect(k,m,t) = 1$ if the m^{th} maintenance task for transmission component *k* begins at time *t*, and 0 otherwise. Then the objective function of our maintenance problem can be expressed as:

$$Max\left\{\sum_{k=1}^{N}\sum_{m=1}^{L_k}\sum_{t=1}^{T}\Delta CR(k,m,t)\times Iselect(k,m,t)\right\} \qquad (1)$$

We assume that each maintenance activity decreases the probability of a particular contingency *k*, which is triggered by failure of the associated equipment. As expressed in Eq. (1), in order to calculate the cumulative risk reduction by maintenance activities, we need to determine the maintenance induced contingency probability reduction, i.e.,

$$\Delta p(k,m) = p_{bm} - p_{am} \qquad (2)$$

where p_{bm} is the equipment failure probability before maintenance, p_{am} is the equipment failure probability after maintenance. Thus, in order to optimize maintenance activities, we need to determine the equipment's time-dependent failure probability based on available condition information.

III. Transformer Failure Rate

This section presents some general concepts in equipment reliability to serve as a reference for the remainder of this paper. Let *T* be a random variable representing the time from when the equipment begins operation at $t = 0$ until a failure occurs. The equipment may be either new or used when it begins operation. In many cases, the time $t = 0$ is after a refurbishment or a failure has been corrected. The uncertainty in the time to failure *T* is described by the distribution function $F(t) = \Pr(T \le t)$, or the probability density function $f(t) = F'(t)$. The probability density function $f(t)$ may also be expressed as:

$$f(t)\Delta t \approx P(t < T \le t + \Delta t) \qquad (3)$$

Hence, $f(t)\Delta t$ is approximately equal to the probability that the equipment will fail in the time interval $(t, t+\Delta t)$. The survivor function, which gives the probability that equipment will not fail up to time *t*, is given by:

$$R(t) = \Pr(T > t) = \int_t^{\infty} f(t)\,dt \qquad (4)$$

The equipment's life distribution is often most effectively characterized by the so-called failure rate, or hazard function, which is the conditional probability of failure. The failure rate function $h(t)$ is expressed as:

$$h(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr[t < T \le t + \Delta t \mid T > t] \qquad (5)$$

If we consider the equipment that has survived the time interval $(0, t)$, i.e. $T > t$, then the probability that the equipment will fail in the time interval $(t, t+\Delta t)$ is approximately $h(t) * \Delta t$. It is only necessary to know one of the functions $h(t)$, $f(t)$, $R(t)$ in order to be able to deduce the other two, as illustrated in Figure 1 [8].

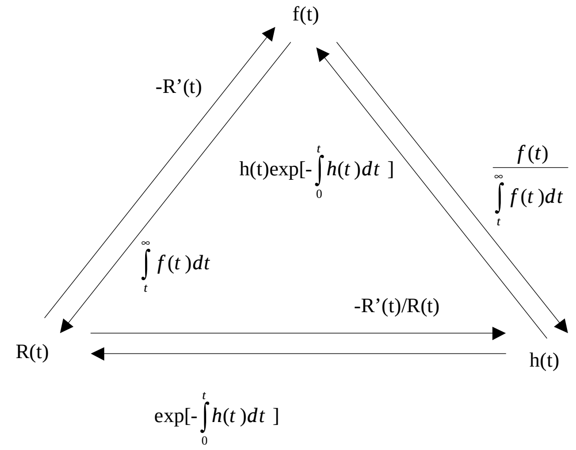

Figure 1: Relationships between h(t), f(t), and R(t)

The Weibull distribution is a widely used distribution to model equipment's time-to-failure. The Weibull probability density function is:

$$f_T(t) = \begin{cases} \dfrac{b\,t^{b-1}}{a^b}\exp\left[-\left(\dfrac{t}{a}\right)^b\right], & t, a, b > 0 \\ 0, & otherwise \end{cases} \qquad (6)$$

β is called the shape parameter because it determines the shape of the distribution. The parameter α is called the scale parameter because it determines the scale. Typically β is between 0.5 and 8.0. As β increases, the mean of the Weibull distribution approaches α and the variance approaches zero. Figure 2 illustrates the Weibull distribution with different shape and scale parameters.

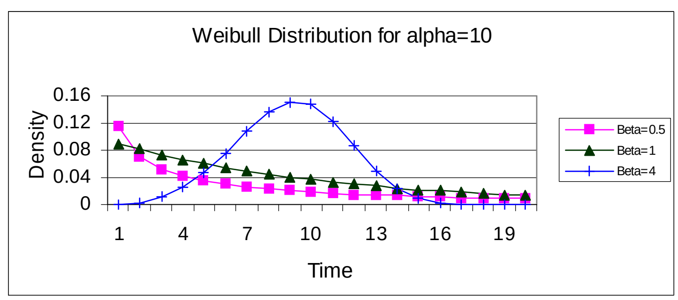

Figure 2: Weibull Distributions

The Weibull hazard function is:

$$h(t) = \frac{b\,t^{b-1}}{a^b}, \quad t > 0 \qquad (7)$$

If $b < 1$, the failure rate is decreasing; if $b = 1$, the failure rate is constant at a value of $1/a$; if $b > 1$, the failure rate is increasing.

IV. Failure Rate Estimation based on Bayesian Analysis

In this section, we describe a Bayesian approach for estimating the power transformer failure rate based on available condition monitoring information. Because power transformer failures tend to be relatively rare events, empirical data for parameter estimation are generally spare. Thus, Bayesian method becomes a natural means to incorporate a wide variety of forms of information in the estimation process.

In the Bayesian framework, the uncertainties in the parameters due to lack of knowledge are expressed via probability distributions. This includes unknown distribution parameters. The Bayesian approach treats the unknown parameter, e.g., α or β in the Weibull characterization of the hazard function, as a random variable. Suppose t is an unknown parameter in our probability model. We first define a distribution, $P(t)$, which generally aim to be as uninformative as possible. $P(t)$ is the prior distribution which represents uncertainty about t based on prior knowledge, e.g. historical information. Then, the posterior distribution of t, given some observations of transformer condition monitoring data, is given by Bayes' Rule:

$$P(t|data) = \frac{P(data|t)P(t)}{P(data)} \tag{8}$$

Here $P(data) = \int P(data|t)P(t)dt$. Suppose the obtained condition monitoring information includes: x_1, x_2, x_3, x_4, etc., which may represent the DGA results, temperatures and other information. Then the conditional distribution $P(data|t)$ takes the form of $P(x_1, x_2, x_3, x_4|t)$, by the product rule of probability, which can be factored as:

$$\begin{aligned} P(x_1,x_2,x_3x_4|t) &= P_{X_4}(x_4|x_1,x_2,x_3,t)\times P_{X_3}(x_3|x_1,x_2,t) \\ &\times P_{X_2}(x_2|x_1,t)\times P_{X_1}(x_1|t) \end{aligned} \tag{9}$$

If x_1, x_2, x_3, x_4 are independently distributed, Eq. (9) can also be written as:

$$P(x_1,x_2,x_3,x_4|t) = P(x_1|t)P(x_2|t)P(x_3|t)P(x_4|t) \tag{10}$$

The resulting posterior distribution is a conditional distribution, conditional upon observing transformer monitoring data. Thus, by using the above Bayesian approach, we can continuously update the transformer failure probability model based on available equipment condition monitoring information, as illustrated in Figure 3.

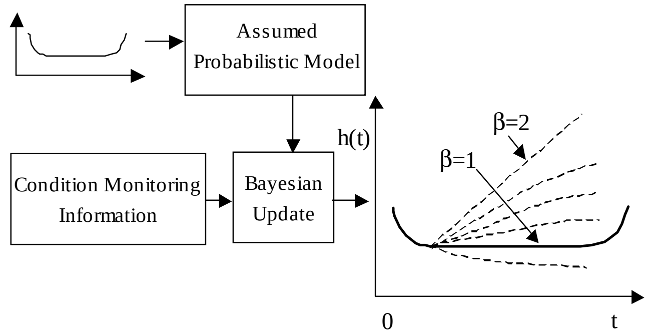

Figure 3: Bayesian Analysis of Transformer Failure Rate

V. An Illustration

We assume the time to failure of a power transformer follows a (mixed) Weibull distribution with scale parameter α and shape parameter β, i.e., $f(t|a,b)$. Since only shape parameter β determines whether the failure rate is increasing or decreasing, for simplicity, we assume scale parameter α is known and only model the uncertainty in the shape parameter β. As described in Eq. (8), although the Bayesian principle is very simple, sometimes it is difficult to apply. First of all, we have to have a prior probability distribution for β. The process of determining the prior can be very involved and sometimes controversial. A useful discussion on this can be found in [9]. Here for illustration purpose, we assume the shape parameter β follows a normal distribution with m and g^2 as its mean and standard deviation, i.e. $b \sim N(m, g^2)$. Here we use available DGA results to update our prior distribution on β. As we all know, the normal family is its own conjugate family. And the use of conjugate family distributions allows some of the unknown parameters to be analytically integrated out, easing the computational burden of the problem. Thus for simplicity, we assume the amount of total dissolved combustible gases, G, also follows normal distribution, i.e. $G \sim N(wb + k_i, w^2 s^2)$, where $i = 1, 2, 3, 4$, corresponds to the conditions indicated by [10]; w is a known parameter; k_i is the average amount of total dissolved combustible gases in condition i which can be referred to [9] if no historical information is available. Then, by the linearity of normal distribution, we can get:

$$Gtemp = \frac{G - k_i}{w} \sim N(b, s^2) \tag{11}$$

Using the above Bayesian update equation (8), we can get the posterior of β, which is still a normal distribution with mean and variance given by:

$$E(b|G) = \frac{g^2}{g^2 + s^2} Gtemp + \frac{s^2}{g^2 + s^2} m \tag{12}$$

$$Var(b|G) = \frac{g^2 s^2}{g^2 + s^2} \tag{13}$$

In this example, we use a set of transformer DGA results obtained from MidAmerican Energy [11], which is given in Appendix A. The data consist of the DGA results for a power transformer over an 8-year period. As stated before, our choice of the normal distribution to analyze this DGA data is for illustration purpose only. The standard approach for analyze such data can be based on Bayesian regression models [12]. The values of these known parameters used in our example are: w =15, k=60, g^2 =2, s^2 =20. And the initial value of m is 1.0. The evolution of means of the posterior distribution of shape parameter β is shown in Figure 4. We specify the scale parameter by taking advantage of empirical transformer failure rate is about 2% per year, i.e. α=50 if β=1. Based on the above expectations of the posterior distribution of the shape parameter, we can estimate power transformer failure rate using Eq. (7), which is shown in Figure 5.

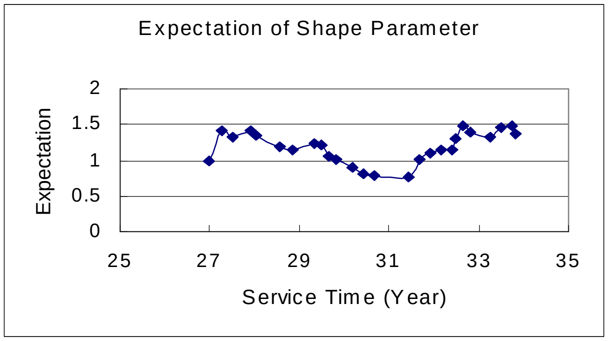

Figure 4: Expectation of Shape Parameter

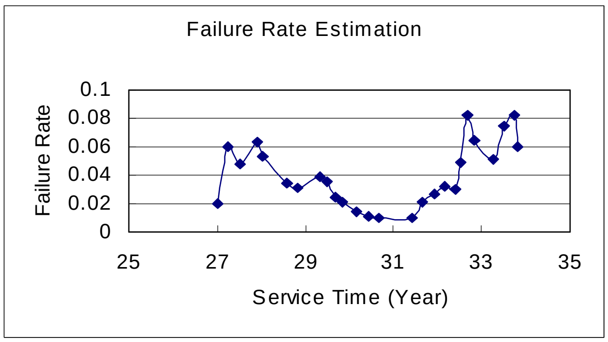

Figure 5: Estimated Transformer Failure Rate

VI. Conclusions

Maintenance is extremely important for transmission equipment health as well as system reliability. The overall objective of our maintenance optimization problem is to maximize the cumulative risk reduction by performing appropriate maintenance tasks. The aim of this paper was to develop a Bayesian approach for estimating the power transformer failure rate based on available condition monitoring information. The Bayesian approach takes advantage of any condition monitoring information that is available through different monitoring techniques. And by tuning the values of known parameters as well as choosing different distributions for unknown parameters, the Bayesian approach is very flexible to estimate failure rate of each individual power transformer with different conditions.

VII. Acknowledgements

This research has been supported by Power Systems Engineering Research Center (PSerc).

VIII. References

[1] CASCADE project: http://www.sis.pitt.edu/~cascade/
[2] http://www.mro.com/corporate/products/maximo.htm
[3] Jaime A. Reinoso-Castillo, "Ontology-driven information extraction and integration from heterogeneous distributed autonomous data sources: A federated query centric approach", M.S. thesis, Department of Computer Science, Iowa State University, Ames, 2002.
[4] EPRI product, Maintenance Management Workstation, http://www.epri.com
[5] ABB Asset Sentry, http://www.abb.com
[6] Yong Jiang, Zhong Zhang, Tim Van Voorhis, J. D. McCalley, "Risk-based Maintenance Optimization for Transmission Equipment", accepted by 35th North American Power Symposium, Rolla, Missouri, October 20-21, 2003.
[7] Ming Ni, McCalley, J. D., Vittal, V., Tayyib, T., "Online Risk-based Security Assessment", Power Systems, IEEE Transactions on, Volume: 18, Issue: 1, Feb 2003, Page(s): 258 –265.
[8] L. Wolstenholme, *Reliability Modeling, A Statistical Approach*, Chapman and Hall, New York, 1999.
[9] Siu, Nathan O., Kelly, Dana L., "Bayesian Parameter Estimation in Probabilistic Risk Assessment", Reliability Engineering and System Safety, ISSN 0951–8320, Volume 62, no. 1–2, pp. 89–116, Oct, 1998.
[10] IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, IEEE Std C57.104-1991.
[11] http://www.midamericanenergy.com/
[12] Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin, *Bayesian Data Analysis*, Chapman &Hall, 1995.

Appendix A

Transformer Oil Test Results

Company: MidAmerican Energy
Manufacturer: Westinghouse
MFG Year: 1968
Primary kV: 161000
Second kV: 69000
Primary KVA Rating: 125,000.00
Equipment Type: AUTO
Phase: 3
Total Gallons: 13,122.00

Sample Date	H_2	CO	CO_2	CH_4	C_2H_2	C_2H_4	C_2H_6	TDCG*
05/12/2003	5	4	729	4	0	7	4	24
04/24/2003	50	0	917	4	0	36	9	99
01/29/2003	64	15	674	3	0	51	13	146
10/15/2002	0	8	1274	7	0	17	8	40
05/13/2002	6	10	771	5	0	8	5	41
03/09/2001	42	58	2301	27	0	14	22	163
01/18/2001	83	15	1062	13	0	27	12	150

12/27/2000	575	129	1227	17	0	52	26	799
09/14/2000	35	10	907	11	0	6	9	71
06/09/2000	45	23	1191	5	3	15	7	98
03/07/2000	55	27	1087	5	0	15	8	110
12/03/1999	90	31	427	5	8	22	9	165
03/23/1999	52	0	917	7	0	3	1	63
12/07/1998	27	13	1818	5	0	12	7	64
09/02/1998	10	10	0	6	0	12	7	45
05/25/1998	11	0	0	6	0	14	8	39
03/03/1998	51	0	388	8	0	1	1	61
01/22/1998	0	1	643	2	0	9	6	38
12/27/1997	0	9	1431	0	0	20	11	71
06/19/1997	27	11	1473	6	1	17	10	98
03/05/1997	0	8	913	7	0	17	10	69
09/05/1996	0	9	1143	0	0	14	8	43
07/11/1996	1	14	1441	6	0	25	10	67
02/26/1996	18	20	1237	7	0	18	10	100
11/16/1995	0	20	1306	8	0	13	10	65
08/16/1995	16	58	2002	7	0	19	10	142

*TDCG: Total dissolved combustible gases. It does not include CO_2, which is non-combustible.